



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/723,344	11/26/2003	Christian Alexander Lang	YOR920030480US1	8235
7590		04/16/2008	EXAMINER	
Ryan, Mason & Lewis, LLP			VAUGHN, GREGORY J	
90 Forest Avenue				
Locust Valley, NY 11560			ART UNIT	PAPER NUMBER
			2178	
			MAIL DATE	DELIVERY MODE
			04/16/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/723,344	Applicant(s) LANG ET AL.
	Examiner GREGORY J. VAUGHN	Art Unit 2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If no period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED. (35 U.S.C. § 133).

Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 31 December 2007.

2a) This action is FINAL. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-4,6-14 and 16-19 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1-4,6-14 and 16-19 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All b) Some * c) None of:

1. Certified copies of the priority documents have been received.
2. Certified copies of the priority documents have been received in Application No. _____.
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)

2) Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
Paper No(s)/Mail Date _____

5) Notice of Informal Patent Application

6) Other: _____

DETAILED ACTION

Action Background

1. This action is responsive to the submission of the Appeal Brief filed on 12/31/2007.
2. Arguments and remarks presented in the Appeal Brief are persuasive, and therefore the finality of the office action dated 10/17/2007 has been withdrawn.
3. Claims 1-4, 6-14 and 16-19 are pending in the case; claims 1, 16, 17, and 18 are independent claims.
4. The examiner's rejection of claims 1-4, 6-14 and 16-19 rejected under 35 U.S.C. 102(e) as being anticipated by Neal et al. US Patent 6,697,799, as described in the office action dated 10/17/2007 is withdrawn in view of the arguments presented in the Appeal Brief filed 12/31/2007, however, new grounds of rejection are made, as described below.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-4, 6-11 and 16-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson et al. US Patent publication 2005/0027664, filed 7/31/2003, published 2/3/2005 (hereinafter Johnson).

7. **Regarding independent claim 1**, Johnson is directed toward an automated document annotation system (see abstract). Johnson discloses receiving an annotation proposed by the user to be associated with a document. Johnson recites: "*In another embodiment or mode of use, learning takes place in the background at the same time that a user annotates a current document and the system provides suggestions to the user in the current document. In embodiments, a user can switch learning modes from iterative to concurrent and vice versa*" (paragraph 27).

Johnson discloses a knowledge base containing allowed annotations in Figure 1 at reference sign 140 (described in paragraph 40 as "database"), where the allowed annotations are stored as "Patterns", "Dictionaries" or "Partially Annotated Text". Johnson discloses matching a user proposed

annotation with an annotation from the database. Johnson recites: "For instance, automatically annotating key information in text as a precursor to indexing can improve search, e.g., if a system annotates the sequence of tokens "International", "Business", "Machines", "Corporation" as a single entity of type "Company" or uses this annotation to further extract and format the information in a simple template or record structure, e.g., [Type: Company, String: "International Business Machines Corporation"], then such information could be used by a subsequent search engine in matching queries" (paragraph 6).

Johnson discloses annotating the document with the allowed annotation. Johnson recites: "The invention generally relates to identifying, demarcating and labeling, i.e., annotating, information in unstructured or semi-structured textual data" (paragraph 2).

Johnson discloses annotating the document as described above. Johnson discloses the system providing a set of matching annotations. Johnson recites: "In this manner, the system and method of the invention produces a final set of one or more annotators to be used by a general annotator-applier on arbitrary text input, which determines specific instances of annotations and in addition, assigns confidence levels indicating the likelihood that annotation instances are correct" (paragraph 27). Johnson discloses the user selecting an annotation from the set of matching annotations. Johnson recites: "This system and method allows the user to provide feedback or supervision in various ways that speed up the learning and annotating process and reduce

the amount of manual effort by, for example, providing for the manipulation of lists of annotated items rather than requiring users to examine tokens in documents and for selective presentation by the system to the user of lists of annotation instances whose confidence levels fall within an (adjustable) confidence range" (paragraph 32). Johnson discloses the system automatically selecting the annotation form the annotation set. Johnson recites: "It should be kept in mind that the training of annotators is completely automatic given the training data, requiring no decisions or actions on the part of the user" (paragraph 43).

Johnson discloses annotating a document from a user supplied annotation that is matched against a database of allowed annotations, as described above. Johnson discloses an annotation selection made by a user or automatically by the system, as described above. Johnson fails to explicitly disclose the case where if the annotation set only contains one item (i.e. only one match to the user proposed annotation) the user need not consider any annotations. However, Johnson is directed toward automation of the annotation of documents, and the goal of Johnson's system is to minimize the effort of the system user. Because Johnson's system can automatically select one item from a set of many, the system could also automatically select one item from a set of one.

Therefore, it would have been obvious, to one of ordinary skill, at the time the invention was made, to provide automatic selection of an item from a set

of one item because "*It would be advantageous to have the text automatically annotated with key information*" (paragraph 9).

8. **Regarding dependent claim 2**, Johnson discloses notifying the user of the results of the annotation matching process. Johnson recites: "*Some of the resulting named entity annotation instances are selectively presented to a user for evaluation and possible correction*" (paragraph 49).
9. **Regarding dependent claim 3**, Johnson discloses storing the user the user proposed annotation/allowed annotation match in Figure 1 at reference sign 141 (shown as "*Dictionaries*").
10. **Regarding dependent claim 4**, Johnson discloses notifying the user that the user proposed annotation matches more than one allowed annotation, as described above.
11. **Regarding dependent claims 6 and 7**, Johnson discloses notifying the user of match results after each matching operation (claim 6) or after a predetermined number of attempted matching operations (claim 7). Johnson recites: "*annotator learning involving iteratively building annotators where at the end of each iteration, the user provides feedback correcting the annotations of the system at that stage, or alternatively a concurrent "walk-through" mode of learning, in which as the user labels data, the learner learns in the background concurrently and makes suggestions to the user*" (paragraph 30).

Art Unit: 2178

12. **Regarding dependent claims 8 and 9**, Johnson discloses maintaining a history buffer of matches (claim 8), and using the history buffer to update the allowed annotations (claim 9). Johnson recites: *"The software modules 120 can access one or more databases 140 in order to read and store required information at various stages of the entire process. The database stores such items as seeds 141, unannotated text 142, annotators 143 including final annotators for exporting and use in runtime applications to annotate message data 144"* (paragraph 40).

13. **Regarding dependent claim 10**, Johnson discloses a history buffer, as described above. Johnson's system allows a user to train a computer to annotate a given document. Johnson discloses an "Interactive Training" mode (starting at paragraph 99) that provides a process to disambiguate matches. Johnson recites: *"In general it is not sufficient to build just a glossary or list of items, rather a system for annotating named entities must have the capability of learning contexts to disambiguate the type of potential entities or class in instances. For instance, "He" could be a pronoun or refer to the chemical element "Helium" and "Madison" might in context refer to a city, a person or some other kind of entity. Therefore, the system and method of the invention cannot simply learn lists of entity mentions, rather it also learns the textual contexts in which particular types of entities occur. By learning the contexts in which named entities of a particular type occur, the system and method can learn to annotate named entities without invoking a specific list or dictionary. The system and method can also learn internal features or characteristics that*

are distinctive of particular classes, e.g., that names of people in English typically have the initial character capitalized, phone numbers consist of digits in various recognizable formats, many addresses have recognizable syntactic characteristics, etc. How to encode this kind of information (internal and contextual linguistic information) into features that can be used as the input to learning algorithms is well understood and common in the field of machine learning. One approach to this is described in detail in Zhang" (paragraph 100).

14. **Regarding dependent claim 11,** Johnson discloses determining a closeness between the user proposed annotation and the allowed annotation in paragraphs 91-98, where the calculation of a matching confidence level is described. The calculated matching confidence level is a measure of the closeness between the user proposed annotation and the allowed annotation.
15. **Regarding independent claims 16, 17 and 18,** the claims are directed toward an apparatus, article of manufacture and method for the method of claim 1, and are rejected using the same rationale.
16. **Regarding dependent claim 19,** Johnson discloses the user proposed annotation matching the allowed annotation, and saving the annotation as described above. Johnson also discloses using the user proposed annotation in a subsequent search, as described above.

17. Claims 12 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson in view of Tunkelang US Patent publication 2003/0120630, filed 12/20/2001, published 6/26/2003.

18. **Regarding dependent claims 12 and 13**, Johnson discloses determining the closeness between terms, as described above. Johnson's determination relies on a confidence level assignment (as described in paragraphs 91-98). Johnson fails to disclose using a term graph to represent the knowledge base, and determining based on computing the distance between terms in the term graph. Tunkelang is directed toward using graphs to determine the distance between search terms. Tunkelang recites: *"The similarity search method and system of the present invention can be used to define and compute the distance between two items in the context of the clustering problem. The clustering problem is often represented in terms of a graph of nodes and edges. The nodes represent the items and the edges connecting nodes have weights that represent the degree of similarity or dissimilarity of the corresponding items"* (paragraph 262).

Therefore, it would have been obvious, to one of ordinary skill in the art, at the time the invention was made, to combine Tunkelang with Johnson in order to provide: *"a similarity search method and system that use an alternate, non-Euclidean approach, are applicable to a variety of data sets, and return results that are meaningful"* (Tunkelang, paragraph 17).

19. Claim 14 is rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson in view of Tunkelang, in further view of Loofbourrow et al., US Patent publication 2004/0006456, filed 7/7/2003, published 1/8/2004.

20. **Regarding dependent claim 14**, Johnson and Tunkelang disclose determining an annotation match by using a term graph, as described above. Johnson and Tunkelang fail to disclose the term graph using a stemming operation. Loofbourrow discloses using a stemming operation associated with search process in Figure 2 at reference sign 16. Loofbourrow recites: *"The processing which takes place within the tokenizer 12 results in the generation of individual word tokens, which are presented to a stemmer 16, where each token is examined and, where appropriate, reduced to its grammatical stem. In the preferred embodiment of the present invention, multi-lingual capabilities are achieved by employing two concepts that facilitate a fast stemming operation"* (paragraph 38).

Therefore, it would have been obvious, to one of ordinary skill in the art, at the time the invention was made, to combine Johnson and Tunkelang with Loofbourrow because: *"This process serves two primary purposes. First, it helps to reduce the size of the index, since all forms of a word are reduced to a single stem, and therefore require only one entry in the index. Second, retrieval is improved, since a query which uses one form of a word will find documents containing all of the different forms"* (Loofbourrow, paragraph 5).

Response to Arguments

21. Applicant's arguments with respect to claims 1-4, 6-14 and 16-19 have been considered but are moot in view of the new ground(s) of rejection, as described above.

Conclusion

22. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Gregory J. Vaughn whose telephone number is (571) 272-4131. The examiner can normally be reached Monday to Friday from 8:00 am to 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Stephen S. Hong can be reached at (571) 272-4124. The fax phone number for the organization where this application or proceeding is assigned is (571) 272-2100.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/Stephen S. Hong/
Supervisory Patent Examiner, Art
Unit 2178

/Gregory J. Vaughn/
Patent Examiner
April 13, 2008